International Environmental Modelling and Software Society (iEMSs)
8th International Congress on Environmental Modelling and Software
Toulouse, France, Sabine Sauvage, José-Miguel Sánchez-Pérez, Andrea Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2016-proceedings

A New GPR Image Analysis Proposal Based on a Multi-agent Approach and Properties of Groups – Towards Automatic Interpretations

David Ayala-Cabrera[ac], Manuel Herrera[b], Joaquín Izquierdo[a], Rafael Pérez-García[a], Olivier Piller[c]
[a]*FluIng-IMM, Universitat Politècnica de València, Cno. de Vera s/n edif. 5C,46022, Valencia, Spain*
[b]*EDEn-ACE Dept., University of Bath, Claverton Down, BA2 7AY Bath, UK*
[c]*Irstea, UR ETBX, Dept. of Water, F-33612 Cestas, France*
e-mail: [a] {daaycab; jizquier; rperez}@upv.es, [b] amhf20@bath.ac.uk, [c] olivier.piller@irstea.fr

Abstract: This paper proposes a new methodology to analyze images from subsoil surveys obtained by using ground penetrating radar (GPR) as a non-destructive method. The aim is to advance towards automatic GPR image interpretations. The main idea is to promote the inclusion of these non-destructive technologies as support in technical management in large-scale systems such as water distribution systems (WDSs). The methodology proposed in this paper uses pre-processing techniques based on a multi-agent approach and analysis of the properties of the obtained groups. In addition, a first classification (via perceptron) of the selected property in the groups is performed. The work is based on GPR studies performed on controlled laboratory conditions; two kinds of test conditions are raised. The first kind pursues capturing objects in GPR images to favor feature extraction thereof, since these tests are used to develop the proposed methodology. The second kind of tests seeks to assess the feasibility of the proposed methodology implementation, while the interaction of the captured objects with other objects and with the environment is evaluated. The results of this study are promising in regard to the selection of areas in GPR images via automatic interpretation that allows the extraction of relevant features for WDS networks. In addition, with the proposed methodology, the use of GPR is promoted by non-highly skillful and expert operators in GPR image interpretation.

***Keywords**:* Image analysis and processing; visualization; multi-agent systems; automatic GPR image interpretation; ground penetrating radar.

1 INTRODUCTION

Currently, the inclusion of some technologies into civil engineering fields, such as water distribution systems (WDSs), is a challenge for researchers and utility managers. Data obtained after the incorporation of certain technologies should improve decision-making processes about the actions to take on the assets of those systems. In this sense, non-destructive methods have shown to be interesting techniques that support network component assessment without affecting the surrounding environment conditions. Ground penetrating radar (GPR) is one of the most reliable tools for obtaining information of buried objects. GPR offers a non-destructive way of exploring the shallow subsurface to detect buried objects such as pipes, cables, ducts and drains.

Once GPR data is acquired, it is usually tackled by both manual and automatic processes. The analysis and interpretation of large volumes of data generated by GPR surveys are extremely difficult and often become a bottleneck that impedes suitable application of the technique. GPR image interpretation is often based on shapes obtained after wave reflections; such as lines and hyperbolae, among others. The shape of the overall response of GPR for buried objects (especially in specific objectives and horizontal cylindrical objects) is a hyperbola (Ahmadi et al. 2015). One of the main strategies used in the interpretation of GPR images is given by the identification and recognition of hyperbolae. One can hardly estimate both the spatial position and the depth of the buried object registered in the GPR images (Janning et al. 2014). Firstly, no knowledge is available about how many

hyperbolae there exist and where they are in a GPR image. In addition, one must confront the added problem of additional reflections that often appear under (or near) the initial reflector. This happens since the GPR signal has a complex waveform and the single reflector may appear as a pair or more reflections. Sometimes this is overlooked, so there is more than one interpretation for the same GPR reflector profiles (Robinson et al. 2013).

Since its commercial release, there have been multiple efforts to overcome one of the major limitations of this technology such as the "human factor" in the interpretation of results, given that the operators have to analyze and interpret a large number of GPR images. This requires a huge effort by the operator and it is not always productive in obtaining results, besides being an inefficient process since it is very time-consuming (Janning et al. 2014; Simi et al. 2008). Speeding up these processes for automating the interpretation is highly desirable. GPR images are complex and difficult to understand, so GPR image interpretation is inherently subjective and depends on the knowledge, skill and experience of the operator (Lu et al. 2014). In this paper, we propose a new methodology to analyze GPR data. The treatment of the data is performed by techniques based on a multi-agent approach and the analysis of the properties in the obtained groups. The properties of the groups are further processed, seeking to extract patterns that allow a basic classification of the existence (or not) of the objects in the evaluated GPR image in a non-subjective way. Additional advantage of the proposed methodology is the possibility of including other properties of the obtained groups on the analysis, in favor of more successful and accurate classifications.

All in all, the main contribution of this work is to provide a new tool for the analysis of GPR images with the aim of automatic interpretations thereof. Thus, understanding of the characteristics of the prospected soil by using non-destructive testing by non-highly qualified personnel is enhanced.

2 OVERAL APPROACH

The signals received in GPR surveys are stored in a matrix, A (radargram, raw data), that is made up of m-vectors, X_j, $j = 1, \dots, n$ (trace), that represent the variation of the soil's electromagnetic properties in terms of depth. Let us represent this matrix by columns $A = [X_1, X_2, \dots X_{n-1}, X_n]$. The length, m, of vectors X_j, is the volume of signal data recorded for each vector, which depends on the characteristics of the equipment used. The sample size is an equipment parameter, generally being sets of 512, 1024, and 2048 samples/trace for commercial equipment. In this study, 512 samples/trace, for 10 ηs/trace, were taken.

The pre-processing of the GPR images used in this document was first proposed by Ayala-Cabrera et al. (2013) and was called *agent race*. The algorithm is based on the game theory and uses the multi-agent paradigm. The input to this algorithm is the A matrix. The n traces generated in the survey are used for the algorithm as parallel tracks for the n agents to run. The race is a test of endurance for the agents. The prize for each agent is a move forward for every effort performed. Efforts are based on wave amplitude values (wav) in each column of the A matrix. The race consists of two phases: a) warming-up and b) competition. The race takes a time $tcomp = tw + tr = m$, where tw is the warming-up time (in this paper $tw = 0$) and tr the competition time. The movements of the agents in tr are conditioned by the reversal of the wave amplitude on the run trace. The race ends when time $tcomp$ has elapsed. The winner is the agent that gets the largest displacement during this time. The output of this process is an array of size $m1 \times n$ where $m1$ = maximum number of movements. By using the peaks (contained in the output) and wav (per column), we rebuild the matrix in the original space ($m \times n$), under the follow conditions: if $\text{peak}_{\text{previous}}(wav) < \text{peak}_{\text{current}}(wav)$; values between these peaks are 1, if $\text{peak}_{\text{previous}}(wav) > \text{peak}_{\text{current}}(wav)$; values between these peaks are -1, and otherwise (peaks, and errors such as clipped waves) values are 0. We denoted the built matrix under the prior conditions as A_{pre}.

The analysis approach by Ayala-Cabrera (2015) suggests that after obtaining the edges through pre-processing of GPR images (e.g. isolines), and by ordering the data as they appear as a function of depth, these data may be represented by groups (families) of functions. The approach to this interpretation is based on the achievement of the matrix Af. This matrix represents the behavior of the data as a collection of functions (families), belonging to three different spaces: the medium (fl), straight horizontal lines (developed in rows); the objects (g), in general constituted by hyperbolas packed in depth; and the noise (r), free individuals. Based on this approach, we proposed the analysis

of matrix A, with the rows used as parallel tracks to run, trying to further eliminate family fl in the results of the pre-processing. Thus, and, in order to not modify the original algorithm, we used $B = A^T$ (T denotes matrix transposition) as input matrix. Once we applied the agent race algorithm by B, we return the results to the original space by the transposition of the matrix A_{pre}; $A_{prerow} = {A_{pre}}^T$. An example of this application is presented in Figure 1.

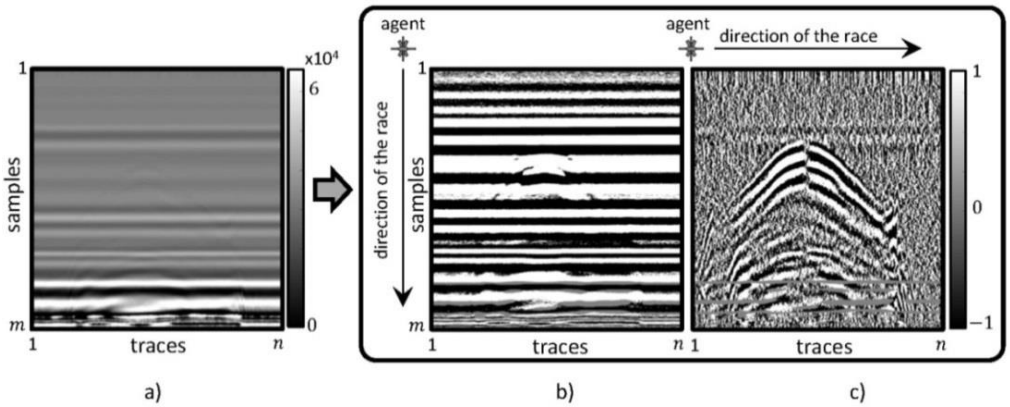

Figure 1. Agent race pre-process. Image: a) A matrix, b) A_{pre} matrix, and c) A_{prerow} matrix.

With the application of the agent race algorithm, using the columns (Figure 1,b) or rows (Figure 1,c), of the input, as parallel tracks for the agents to run, in both cases the visualization of the GPR image is substantially improved with reference to the input (Figure 1,a). In addition, we would like to mention that the applied approach based on families of functions is verified in a visual way with the obtained images. First, in Figure 1,b fl it is the predominant family, and secondly, in Figure 1,c family g (target of this work) it is spotlight and r it is the predominant family. Thus, based on the type of the pre-processing applied in Figure 2,c, we try to extract information of the objects (if there exists) from GPR images (Section 3).

3 AREA DELINEATION, AND GROUP EXTRACTION AND TREATMENT

Once A_{prerow} matrix is obtained, in this section we proposed the division of this matrix into two binary matrices denoted as U and D. The conditions to realize this data division are: $U = \{1, \text{if } A_{prerow} = 1 \vee A_{prerow} = 0; \text{otherwise } 0\}$ and $D = \{1, \text{if } A_{prerow} = -1 \vee A_{prerow} = 0; \text{otherwise } 0\}$. Values 0 in A_{prerow} matrix represent the peaks of the traces and agent interpretation mistakes; we have decided to include (as a pivot) the value 0 in both conditions to build the U and D matrices. Edge detection and group extraction was performed for U and D matrices using the Moore-Neighbor tracing algorithm modified with the Jacob's stopping criteria (Gonzales et al. 2004), implemented in Matlab's *bwboundaries* function. In this study, we focus on two of the properties of the form, V, by the extracted groups from U and D matrices; these properties are: area and centroid; $V(area, centroid(Xc, Yc))$. In addition, the analysis for the centroid is only proposed by Xc; being reduced our analysis to $V(area, Xc)$. The returned order of the properties by the used algorithm is given as a vector $P \times 1$, where P is the length and the number of extracted groups from each of the binary images analyzed (U and D). This order is used in our analysis and we denote it as "data appearance order". An example of the properties of form by U and D, is shown in Figure 2.

The images of matrix A_{prerow} (Figure 2,a) show that the contained object in the image is mainly composed by groups with larger area, in comparison with the groups that are no part of the object. This is reflected in the areas obtained for the different groups in both matrices U and D (Figure 2,b). This way of presenting the properties of the groups, shows a demarcated set of groups protruding of the rest of the groups on the area property (for both matrices – U and D). These protruding groups belong to the object contained in the image. Thus, in this study we call "evident property", a property which the following restriction: $\{Fc(V) = mc \cdot Fc(V) + bc \mid mc = 0,\ Fc(V) > bc :\ \text{true},\ Fc(V) \leq bc :\ \text{false}\}$. It is true for groups belonging to the object and false for groups that do not belong. This evident property provides extraction information about interest groups in the own object and about the other properties to be evaluated; such as Xc (Figure 2,c). $Fc(area)$, presented in Figure 2,b, was obtained from the maximum value of the areas by three tests performed without object inside the image (test

are presented in Section 4). Maximum values of the areas for tests without object, in samples×traces, were: [(test 49; area=2097), (test 50; area=3370), (test 51; area=2236)], with average 2568, and we used this value as $Fc(area)$ as a start point. The membership of the groups to the object (true) was an activity performed by an expert.

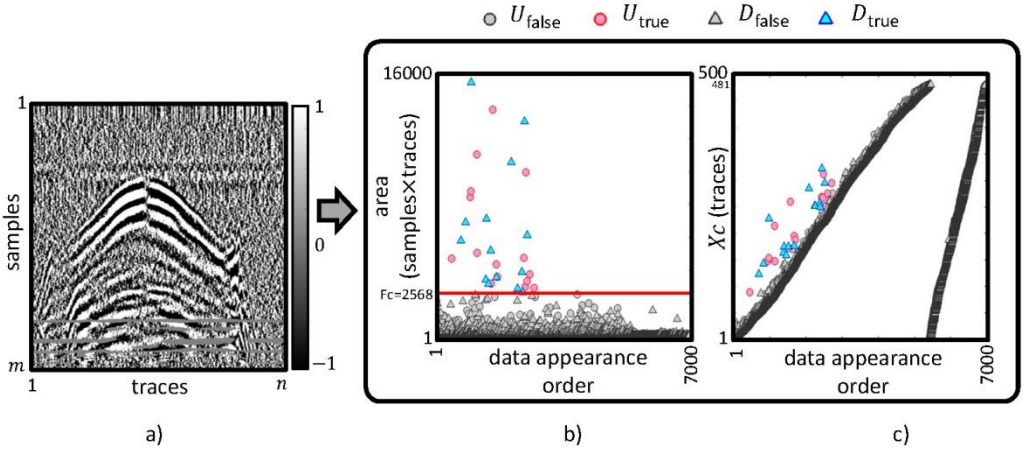

Figure 2. Selected properties – Classification based on area property. a) Image of A_{prerow} matrix, property: b) area, and c) Xc.

Proposed transformation of the space through the Xc property. It is observed in Figure 2,c, that the classified groups as true (in both matrices – U and D), by $Fc(area)$, protrude as two inclined straights just formed by the groups classified as false. Based on this observation and due to its presence in all the tests performed in this work, we try to find a suitable transformation of this data to provide the information of Xc as an evident property. In this sense, our proposal consists in the processing data from Xc (by U and D) using the race agent algorithm. The obtained peaks from this process are used as vertices and from each vertex we rebuild a vector, termed $Xmul$, with the same length as the original Xc (see Figure 3,a). The difference between $\{Xc - Xmul\}$ provides us with a transformation of the space that allows classification of the groups as an evident property, in the same way as the area property (see Figure 3,b).

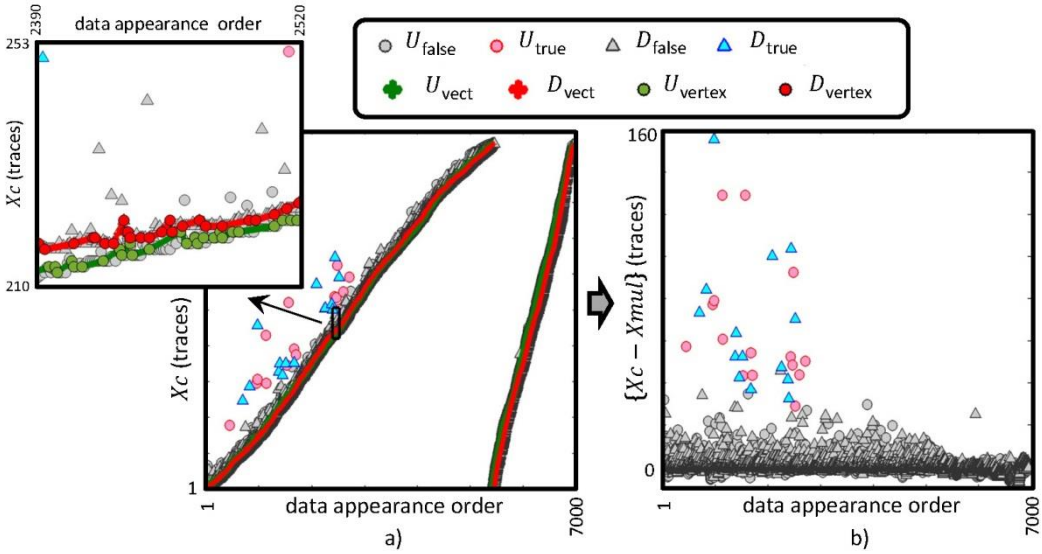

Figure 3. Proposed change of the space by Xc property. a) Xc, and b) $\{Xc - Xmul\}$.

We can note that the change of the space of Xc using the race agent algorithm, enable us more clearly differentiation between true and false classification for the groups (Figure 3). With this transformation it is possible the construction of $Fc(\{Xc - Xmul\})$ in an easy way. In the next section, the use of this transformation of the Xc property, via perceptron and laboratory test, is proposed.

4 COMPARISON BASED ON – Xc AND $\{Xc - Xmul\}$ VIA PERCEPTRON

In this section we have performed laboratory tests to evaluate the feasibility of classification of the groups by the Xc property and its change to space $\{Xc - Xmul\}$. In this regard, we have proposed captures of the object with GPR that contain only the object, in air and with as less interferences as possible. This seeks to get the maximum possible information that can provide the object to be recorded in GPR images. The set of tests used for the current comparison corresponds to 51 tests with different materials and different depths, under unique capture conditions. Objects' (pipes in this paper) materials chosen for the performed tests are: Polyvinyl chloride (PVC), polyethylene (PE), cast iron (CI), and asbestos cement (AC), all of them with a nominal diameter of 10.0 cm. The used depths (in reference to the object point with lower depth) are: 10.5 cm, 23.0 cm, 37.0 cm, and 62.0 cm. Three captures were performed for each proposed depth and material (in total 48 test), and three additional tests with no object (NOB) were taken.

The GPR equipment used in each survey was a commercial monostatic antenna with a central frequency of 1.5 GHz. The parameters of the equipment were 120 trace/s, 2048 samples/trace and 10 ηs/2048 samples.

Perceptrons are a natural choice for branch prediction because they can be efficiently implemented in hardware. One benefit of perceptrons is that by examining their weights, i.e., the correlations that they learn, it is easy to understand the decisions they make. In contrast, a criticism of many neural networks is that it is difficult or impossible to determine exactly how the neural network is making its decision; the perceptron's decision-making process is easy to understand as the result of a simple mathematical formula (Collins et al. 2001). Thus, at this point we use the bounties offered by the analysis via perceptron and looking for verifying the effectiveness of the change of the space $\{Xc - Xmul\}$. We seek that these properties serve as support for the training of intelligent classification systems. The perceptron used in this work responds to (1),

$$Clas_{PR} = w_0 + \sum_{i=1}^{npro} V_i w_i \ , \tag{1}$$

where: $Clas_{PR}$ is the prediction of the classification obtained by the perceptron, $w_{0\ldots npro}$ is the vector of weights, V_i (predictors) are the vectors of the group properties in which the analysis is based. Our main objective is to compare the effectiveness of the change of the space, $V(Xc)$ with $V(\{Xc - Xmul\})$, $npro = 1$. The $error$ is computed as

$$error \equiv \sum_{k=1}^{np} \left[e^{(k)}\right]^2 = \sum_{k=1}^{np} \left[Clas_{PR}^{(k)} - Clas^{(k)}\right]^2 \ , \tag{2}$$

where: $Clas$ is the interpretation vector (true, false) given by the expert based on $Fc(area)$; here we used a value of 1 for true and 0 for false, np is the total number of groups used for the classification, $k = 1, \ldots, np$.

In Figure 4, the results of the perceptron application error for classification of the groups obtained in the 51 test outlined in this section are presented. In this image, two predictors, Xc y $\{Xc - Xmul\}$, trained separetely may be distinguished. Also, and additional condition may be observed, noted by prefixes ind and all, explaining if the groups of the various tests were used to individually train each perceptron (ind), or if the groups obtained for all the tests were used to train just a perceptron (all).

In Figure 4,a we can see that for the error curves no clear relationships exist with the material of the object sought. However, a light inflection is observed in the error curves when reaching larger depths, for all the materials. Additionally, accumulated error curves show that the error does not increase for images with no objects (NOB). For $\{Xc - Xmul\}_{ind}$ the results show better adaptation to data in comparison with $\{Xc\}_{ind}$. This clearly favors predictions closer to the performed classifications, which is reflected by the slope of the accumulated error. We have also obtained that the training data under the individually modified space, ($\{Xc - Xmul\}_{ind}$), the error increase derived from a global classification, ($\{Xc - Xmul\}_{all}$), is very low. In both cases, the slope of the accumulated error is

considerably smaller than the obtained in the original space $\{Xc\}_{ind}$. Moreover, the test errors (Figure 4,b) show peak reduction for $\{Xc - Xmul\}_{ind}$) and $\{Xc - Xmul\}_{all}$), in comparison with $\{Xc\}_{ind}$. In this last case, a clear difference may be observed regarding the minimum depth of the object. This reduction favors the detection of this type of objects, at various depths.

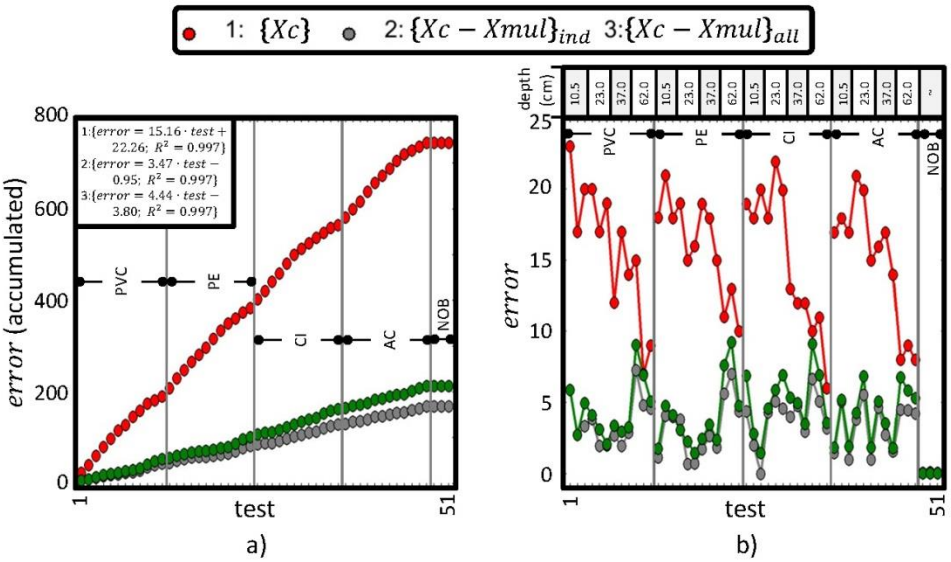

Figure 4. Error curves – Classification via perceptron based in Xc and $\{Xc - Xmul\}$. a) Accumulated $error$, and b) $error$.

Specifically, a correct classification of 370532 groups for 51 images was performed; 744 of those groups were classified as true. The main idea is that true values remain in the object space, while the *falses* are confined to the no-object space. The space division is conditioned by Fc, and the further the data in those spaces, the better the classification. The results for the change of space $\{Xc - Xmul\}$ and the results for the classification obtained with the perceptron for all the data $\{Xc - Xmul\}_{all}$ are presented in Figure 5. In this figure, Fc takes the maximum value obtained in the NOB groups.

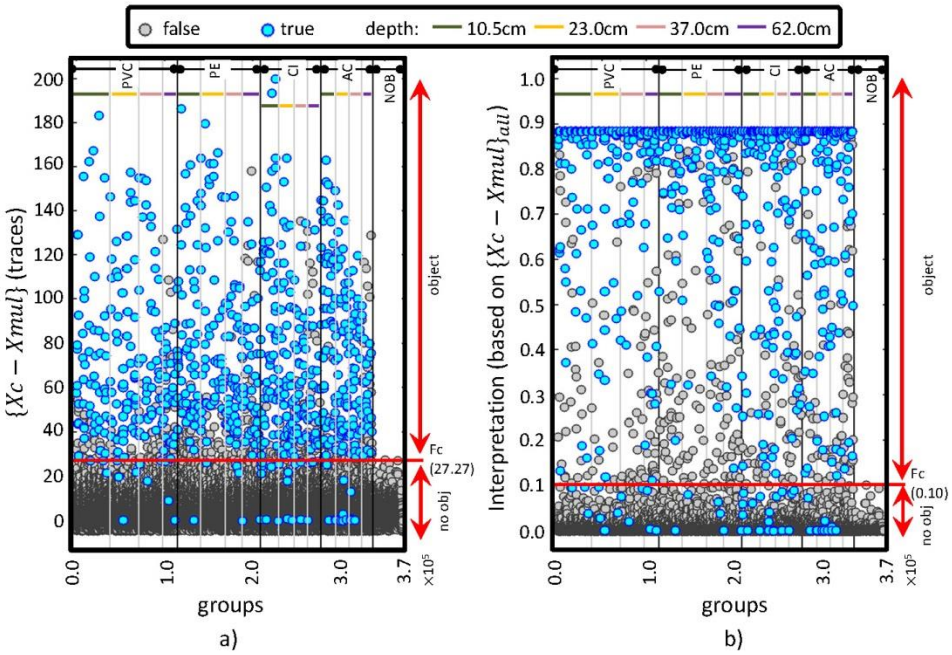

Figure 5. Analysis of groups for $\{Xc - Xmul\}$. Data classification: a) based in the relation $\{Xc - Xmul\}$, and b) via perceptron based in the $\{Xc - Xmul\}_{all}$.

Figure 5,a shows how, by transforming the space, data presents, for the entire data set, a clear classification into objects and no-objects. It is also observed that data classification $\{Xc - Xmul\}_{all}$ via perceptron empowers this natural separation due to the increase in the distance true-false,

corresponding to object and no-object, respectively. As an example, in Figure 6, the groups for test 1 (material PVC, depth 10.5 cm, groups [1, 13883]) are retrieved into the original spaced with their corresponding classification values.

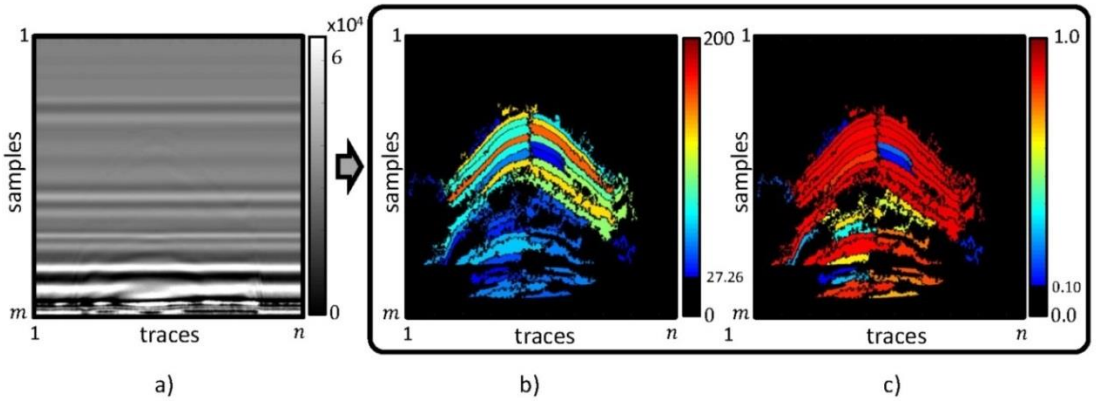

Figure 6. Test 1 – PVC. a) Raw image, Image – Data classification: b) based in the $\{Xc - Xmul\}$ relation, and b) via perceptron based in the $\{Xc - Xmul\}_{all}$ relation.

In Figure 6, it can be observed that the difference between the groups obtained from both classifications (parts b and c) is almost negligible. However, the tendency to disappear from the image is higher in part (b) than in part (c). This enhances the importance of generating classifications more clearly separating the data. In both cases, the visualization improvement of the characteristics of the objects with regard to the raw images (Figure 6,a) is evident.

5 CASE STUDIES

In this section, we implement the interpretation via perceptron based in the relation $\{X - Xmul\}_{all}$. We have to note that training has been performed under conditions of objects with data unrelated to other objects. However, we also increase the difficulty of the study, by evaluating objects in more complex (lab) environments, where interactions among objects are taken into account. To this purpose, we have buried 5 pipes of nominal diameter 100 mm (except for PE2 with nominal diameter 50 mm), very close among them in a lab tank, to check if discrimination is feasible. This time all the pipes were buried simultaneously, at various elevations and positions. With this test we try to approximate the real conditions that may be expected in the field, were various pipes may be buried very close ones to the others. The classification results obtained via perceptron and the schematic layout of the pipe location into the tank are presented in Figure 7. We have to mention that captures were performed three times.

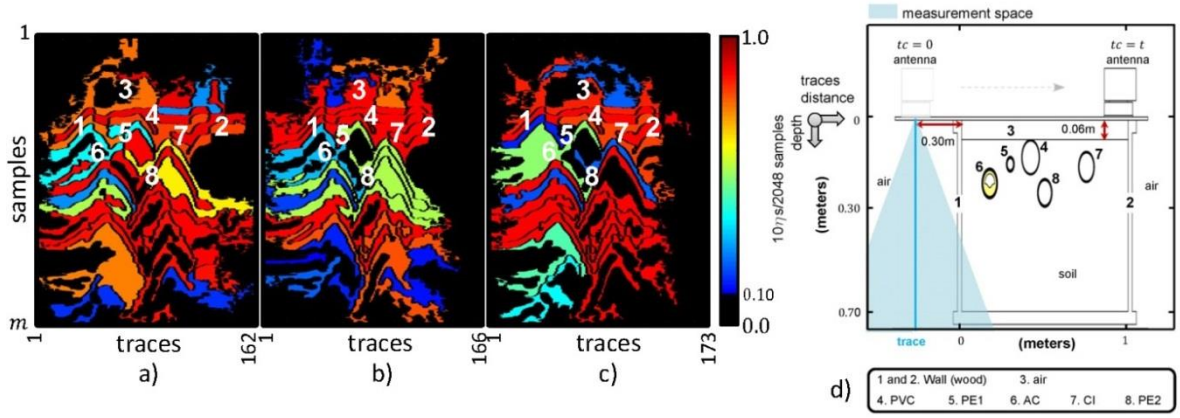

Figure 7. Application of perceptron in complex images – various objects, various materials, short distance among objects. a), b), c), GPR images obtained after perceptron application, and d) schematic configuration.

When comparing the three tests (parts a, b, c) with their respective schematic configuration, we can observe that, by identifying the objects IDs, the appearance sequence is coherent. Analogously, we

observe that the obtained perceptron enables us to go further in the interpretation, thus reducing the analysis difficulty regardless pipe material and size. This makes objects visible, which is especially outstanding for the case specifically problematic materials such as plastics.

6 CONCLUSIONS

This document proposes a new methodology for analyzing GPR images. This methodology is based on techniques using a multi-agent approach, and analysis of groups. As a result of this work it is obtained that the inclusion of new variables not usually explored, as the group form properties, specifically the centroid Xc components, favors clear progress towards automatic interpretation of GPR data. In addition, the methodology proposed in this paper allows the extraction of patterns in an easy, simple and fast way. The patterns obtained enables us to classify the sampled image space, according to the existence of the object(s) in the image.

REFERENCES

Ahmadi, R., Fathianpour, N., Norouzi, G.H., 2015. Detecting physical and geometrical parameters of some common geotechnical targets through their effects on GPR responses. Arabian Journal Geosciences. 8 (7), 4843–4854.

Ayala-Cabrera, D., 2015. Characterization of components of water supply systems from GPR images and tools of intelligent data analysis, Ph.D. thesis, Universitat Politècnica de València, Valencia (Spain).

Ayala-Cabrera, D., Izquierdo, J., Montalvo, I., Pérez-García, R., 2013. Water supply system component evaluation from GPR radargrams using a multi-agent approach. Mathematical Computer Modelling. 57 (7–8), 1927–1932.

Gonzales, R.C., Woods, R.E., Eddins, S.L., 2004. Digital image processing using MatLab, Pearson Prentice Hall, Upper Saddle River, NJ (USA).

Collins, J.D., Tullsen, D.M., Wang, H., Shen, J.P., 2001. Dynamic speculative precomputation. Proceedings of the 34th Annual ACM/IEEE International Symposium on Microarchitecture, Austin, TX, USA.

Janning, R., Busche, A., Horváth, T., Schmidt-Thieme, L., 2014. Buried pipe localization using an iterative geometric clustering on GPR data. Artificial Intelligence Review. 42 (3), 403–425.

Lu, Q., Pu, J., Liu, Z. 2014. Feature extraction and automatic material classification of underground objects from ground penetrating radar data. Journal of Electrical and Computer Engineering, 2014 ID 347307, 1–10.

Robinson, M., Bristow, C., McKinley, J., Ruffell, A., 2013. Ground penetrating radar, Part 1, Sec: 5.5. In Clarke, L.E & Nield, J.M. (Eds.) Geomorphological Techniques (online Edition). British Society for Geomorphology, London (UK), pp. 1–26.

Simi, A., Bracciali, S., Manacorda, G. (2008). Hough transform based automatic pipe detection for array GPR: algorithm development and on-site test. Proceeding of the RADAR'8, IEEE Radar Conference, Rome, Italy.